

Section 508 Compliant ☐ Yes ☒ No



Data Processing
UCD TI #351B, Version 1.4
January 7, 2025
Page 1 of 33

UCD IMPROVE Technical Information #351B

Data Processing

*Interagency Monitoring of Protected Visual Environments
Air Quality Research Center
University of California, Davis*

*January 7, 2025
Version 1.4*

Prepared By:	<div>Signed by:  19F6B63B1B17443...</div>	Date: 1/10/2025
Reviewed By:	<div>DocuSigned by:  2B7AC07297404D6...</div>	Date: 1/10/2025
Approved By:	<div>DocuSigned by:  0A10CECE79B0452...</div>	Date: 1/10/2025

UCDAVIS
AIR QUALITY RESEARCH CENTER

Section 508 Compliant ☐ Yes ☒ NoData Processing
UCD TI #351B, Version 1.4
January 7, 2025
Page 2 of 33**DOCUMENT HISTORY**

Revision	Release Date	Initials	Section/s Modified	Brief Description of Modifications
1.0	03/14/2022	SRS	All	Previously anthologized version separated into individual TIs.
1.1	11/10/2022	DEY, ITS	9	Corrected flow processing examples. Corrected Pre- and post-weight units.
1.2	01/04/2023	DEY, ITS	9.4	Updated analytical MDLs and fractional uncertainties.
1.3	03/03/2023	DEY, ITS	9.3	Updated information on volume calculations for old controllers.
1.4	01/07/2025	ITS, AMM	3, 9, 11.4	Updated definitions, added flow processing and status information and updated analytical MDLs and fractional uncertainties. 11.4 nonconformance and deviation section removed (redundant).

TABLE OF CONTENTS

1. Purpose and Applicability	5
2. Summary of the Method	5
3. Definitions.....	5
4. Health and Safety Warnings	6
5. Cautions	6
6. Interferences.....	6
7. Personnel Qualifications	6
8. Equipment and Supplies	6
9. Procedural Steps.....	7
9.1 Units.....	10
9.2 Artifacts.....	10
9.3 Volume.....	12
9.4 Concentration, Uncertainty, and Method Detection Limit	14
9.4.1 PM _{2.5} and PM ₁₀ Mass (1A and 4D Modules).....	16
9.4.2 Ions (2B Module).....	18
9.4.3 Carbon (3C Module).....	19
9.4.4 Elements (1A Module).....	22
9.4.5 Laser Absorption (1A Module).....	26
9.5 Equations of Composite Variables.....	28
9.5.1 Sulfate (3× sulfur from XRF) and Ammonium Sulfate (NH ₄ SO ₄)	28
9.5.2 Ammonium Nitrate (NH ₄ NO ₃).....	29
9.5.3 Soil	29
9.5.4 Non-Soil Potassium (K ₂ NON)	30
9.5.5 Organic Carbon by Mass (OMC).....	30
9.5.6 Black Carbon	30
9.5.7 Reconstructed Mass Using Carbon Measurements (RCMC)	31
9.5.8 Reconstructed Fine Mass (RCMN).....	31
10. Data Processing Code	32
11. Quality Assurance and Quality Control.....	32
11.1 Code Development.....	32

Section 508 Compliant ☐ Yes ☒ No

11.2 Bug Reporting 33

11.3 Data Validation 33

12. References 33

LIST OF FIGURES

Figure 1. Data processing flow chart. 7

LIST OF TABLES

Table 1. Units for data delivered to the CIRA, AQS, and UCD CIA databases..... 10

Table 2. Universal flow constants for the V4 controllers. 14

Table 3. Analytical method detection limits (MDL) in µg/filter for mass. 17

Table 4. Fractional uncertainty for the mass..... 17

Table 5. Analytical method detection limits (MDL) in µg/filter for the ions species. 19

Table 6. Fractional uncertainty for ions. 19

Table 7. Analytical method detection limits (MDL) in µg/filter for the carbon species. 21

Table 8. Fractional uncertainty for the carbon species. 22

Table 9. Analytical method detection limits (MDL) in µg/cm2 for the elemental species. 24

Table 10. Fractional uncertainty for the elemental species..... 25

Table 11. Analytical method detection limit for τ₆₃₃ (unitless). 27

Table 12. Fractional uncertainty for the laser absorption data..... 28

1. PURPOSE AND APPLICABILITY

The purpose of this technical information (TI) is to provide information on the steps for processing the sampling and analytical data from the Interagency Monitoring of Protected Visual Environments (IMPROVE) network. The raw operational information from field sampling is combined with laboratory analytical results to generate concentrations, uncertainties, and method detection limits.

2. SUMMARY OF THE METHOD

The University of California, Davis (UCD) analyst will use functions in the *crocker* software package to calculate final results and post them to the UCD IMPROVE database. The analyst will also review any output messages for errors.

3. DEFINITIONS

- **AQRC:** Air Quality Research Center.
- **AQS:** EPA's Air Quality System database.
- **CSN and IMPROVE Archive (CIA) Database:** A database of the complete record of CSN and IMPROVE data coupled with a web-based visualization and analysis tool.
- **Chemical Speciation Network (CSN):** EPA's PM_{2.5} sampling network, with sites located principally in urban areas.
- **CIRA:** Cooperative Institute for Research in the Atmosphere.
- **crocker:** A custom software package in the R language that contains the data processing code used to produce, check, and post the final results.
- **Energy Dispersive X-Ray Fluorescence (EDXRF):** An analytical technique used to determine the concentration of elements.
- **Federal Land Manager Environmental Database (FED):** a database of environmental data managed by the Cooperative Institute for Research in the Atmosphere (CIRA)
- **Hybrid Integrating Plate/Sphere (HIPS):** An analytical technique for optical absorption.
- **Ion Chromatography (IC):** An analytical technique used to determine the concentration of ions.
- **Interagency Monitoring of Protected Visual Environments (IMPROVE):** Federal PM_{2.5} and PM₁₀ sampling network directed by the National Park Service, with sites located principally in remote rural areas.
- **IMPROVE database:** A SQL Server database that is the central warehouse of IMPROVE preliminary and final data at UCD.
- **Method Detection Limit (MDL):** A lower limit of detection specific to method of analysis and reported parameter.

- **NPS:** National Park Service.
- **PM:** Particulate Matter. PM_{2.5} is particulate matter with diameters 2.5 micrometers (µm) and smaller. PM₁₀ is particulate matter with diameters 10 µm or smaller.
- **SOP:** Standard Operating Procedure.
- **SQL:** database management system used by AQRC.
- **Thermal Optical Analysis (TOA):** An analytical technique used to determine the concentration of carbon, also referred to as TOR (Thermal Optical Reflectance) and TOT (Thermal Optical Transmittance).
- **TI:** Technical Information; subset document paired to an SOP.
- **UCD:** University of CA—Davis.
- **Sample Handling laboratory (SHL):** The filter handling laboratory for IMPROVE at UCD.
- **Field Group:** The group in charge of samplers and operation of IMPROVE network.
- **UCD IMPROVE Data Management Website:** A web-based application developed with .NET framework to interact with the IMPROVE database. Also referred to as IMPROVE web app.

4. HEALTH AND SAFETY WARNINGS

Not applicable.

5. CAUTIONS

Not applicable.

6. INTERFERENCES

Not applicable.

7. PERSONNEL QUALIFICATIONS

The UCD Air Quality Research Center (AQRC) Data & Reporting Group staff assigned to tasks described in this document have advanced training in database programming and database management.

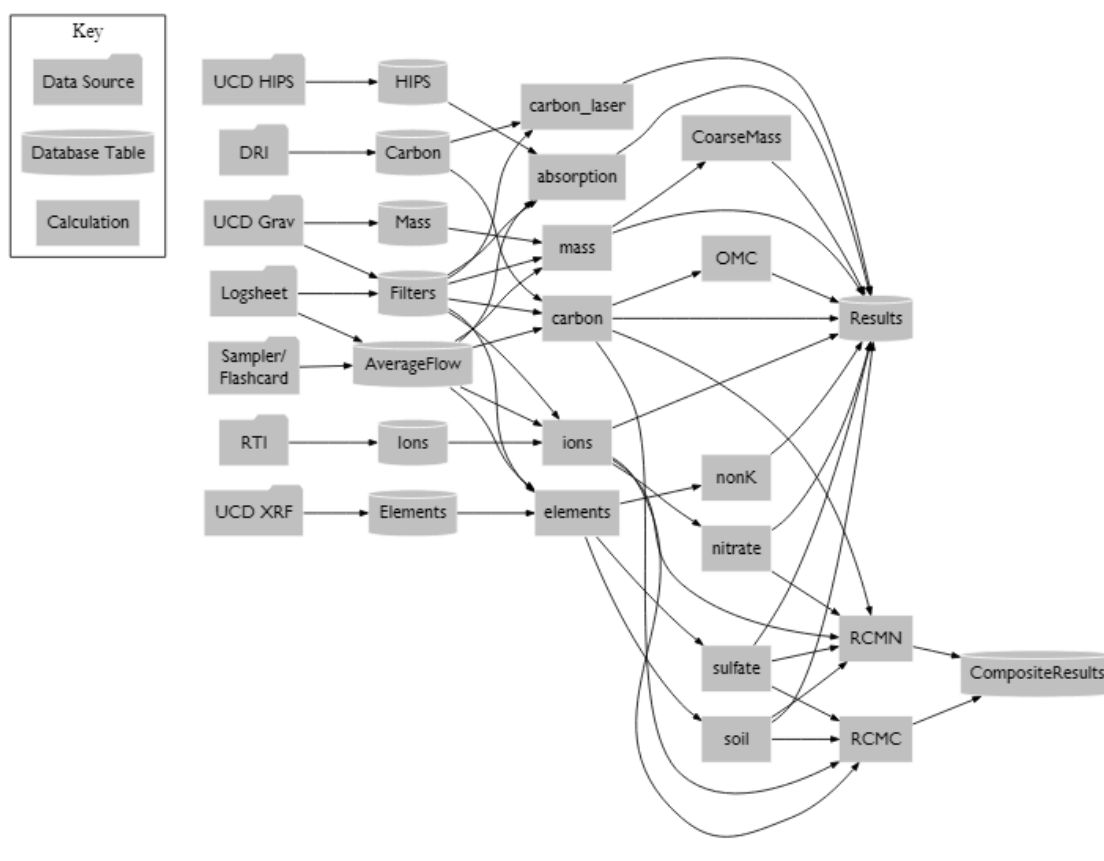
8. EQUIPMENT AND SUPPLIES

The hardware and software used for IMPROVE data ingest are described in *UCD IMPROVE SOP #351: Data Processing & Validation*.

9. PROCEDURAL STEPS

Data processing for IMPROVE consists of reducing and combining data from the sampling and analytical laboratories to calculate concentrations, uncertainty estimates, and method detection limits (MDLs). Figure 1 shows a flow chart for IMPROVE data processing.

Figure 1. Data processing flow chart.



Calculation of concentrations and associated uncertainties and MDLs are performed within the *crocker* R package. The flow rate calculations can be performed in the *crocker* R package or in the UCD IMPROVE database. Flow rate calculations are performed before calculating concentrations to ensure the most up-to-date flow data are used.

Flow data are processed using a stored SQL procedure to derive the daily average flow rate and elapsed time (ET). The flow processing code automatically assigns non-normal flow status flags to the samples with flow rates that deviate from the nominal values.

Samples with nominal flows will receive a normal status. The criteria for assigning flags can be found in Tables 1 and 2 of the UCD IMPROVE TI #351E: Flow_Validation.

The stored SQL procedure is shown below. The first six lines of the SQL query below state the variables to process flows with. In general, the start and end dates are declared to cover the month(s) of data being processed, and the samplername is left blank to process flow data for the entire network. The flow processing can be performed on a single site, date, or even filter ID by declaring the appropriate values.

```

DECLARE @RC int
DECLARE @iStartDate datetime = 'mm/dd/yyyy'
DECLARE @iEndDate datetime = 'mm/dd/yyyy'
DECLARE @iSamplerName NVARCHAR(50) = NULL
DECLARE @iFilterId BIGINT = NULL
DECLARE @Debug bit = 1

EXECUTE @RC = [Improve_2.1].[sampler].[spFilterAverageFlowRates]
    @iStartDate
    ,@iEndDate
    ,@iSamplerName
    ,@iFilterId
    ,@Debug
GO

```

A message will appear in the console upon successful completion of flow processing, displaying the date and time of completion.

If the execution code fails, evaluate the warning message and work with the Software & Analysis Group and/or Sample Handling Laboratory to identify the issue and resolve it. One common issue leading to failure in flow processing is the incorrect assignment of flow source codes. The PM_{2.5} and the PM₁₀ modules have different flow source codes; an incorrect assignment will cause the flow processing procedure to fail.

The flow data can also be processed in the R environment by the *improve_process_flow* function from the *crocker* package. This function uses the main stored procedure (spFilterAverageFlowRates; detailed above) in the IMPROVE database for processing 24-hr average flow rates. To perform flow data processing in the R environment, open an R environment (such as RStudio) and run the following command:

```

[flow] <- crocker::improve_process_flow(start_date = ['YYYY-MM-DD'],
    end_date = ['YYYY-MM-DD'], sampler = [NULL], filter_id = [NULL], server =
    "production")

```


The start date, end date, and server information are required fields like the SQL procedure. Similar to the SQL procedure, the flow processing function can be performed on a single site, date, or even filter ID by declaring the appropriate values.

For processing flow for the entire network, the following code can be used:

```
[flow] <- crocker::improve_process_flow(start_date = ['YYYY-MM-DD'],
end_date= ['YYYY-MM-DD'], server = "production")
```

If successful, the code will return a value of zero. If another value is returned, try to run the code directly in SQL in debugging mode to investigate the error messages and inform the software group.

To calculate a single month of concentration, uncertainty, and MDL for all IMPROVE parameters, including reported and non-reported diagnostic parameters, the following command is run the in the R environment:

```
[month_data] <- crocker::improve_calculate_and_post([YYYY], [MM], skip =
[NULL], server = 'production', AnalysisQcCode = 1, comment = ['Initial
Posting'], replacingId = [NULL], replacingQcCode = [NULL])
```

This command calculates concentrations, uncertainties, and MDLs for all measured and derived parameters for the year (*[YYYY]*) and month (*[MM]*), using all data from the production database (when *server* = 'production'), and appends the processed data to the analysis.Results and analysis.CompositeResults table in the UCD IMPROVE production database (Improve_2.1) as an analysis set. It also inserts a record into the analysis.ResultsSets table that provides summary information for this set, including the *comment* and *AnalysisQcCode*. Routine data uses *AnalysisQcCode* = 1.

The *skip* argument gives the flexibility to skip one or more categories of analysis in processing. All analyses except those listed in the *skip* argument will be posted to the database. Any parameters that are derived from the skipped parameters are also skipped. A typical command using the *skip* argument can be as follows:

```
November_data <- crocker::improve_calculate_and_post(2020, 11, server =
"production ", skip = "optical")
```

If the user wants to skip multiple analysis categories, the categories can be grouped in the *skip* argument, as shown below. Here is a complete list of options:

"mass", "elements", "ions", "carbon", "optical", "flow", "composite"

```
November_data <- crocker::improve_calculate_and_post(2020, 11, server =
"production", skip = c("optical", "elements"))
```

After level 2 validation, the data is reprocessed to capture changes. The flow data is then reprocessed, and the `improve_calculate_and_post` function is executed again. The updated dataset is uploaded to the database. While reposting the data, to ensure version control and maintain data integrity, the following actions must be taken:

- Add comment to describe the new dataset.
- Change the analysis QC code of the previously posted dataset(s) by including the data set ID of the previous posting (*replacingId*) and the analysis QC code (*replacingQcCode*) that should be associated with that data set.

9.1 Units

Table 1 lists the data types, parameters, and units for all data delivered to the CIRA, AQS, and UCD CIA databases (see *UCD IMPROVE TI #351D: Data Delivery*). For mass, ions, carbon, elements, and light absorption, the units listed are also used for uncertainty and MDL. NA indicates that the data type is not reported to the corresponding database.

Table 1. Units for data delivered to the CIRA, AQS, and UCD CIA databases.

Data type	Parameter	CIRA unit	AQS unit	UCD CIA unit
Flow Rate	Flow	L/min	NA	NA
Elapsed Time	ET	min	NA	NA
Gravimetric mass	PM2.5, PM10	ng/m ³	µg/m ³	µg/m ³
Ions	Cl-, NO ₂ , NO ₃ , SO ₄	ng/m ³	µg/m ³	µg/m ³
Carbon	OC1, OC2, OC3, OC4, OC, OPTR, EC1, EC2, EC3, EC	ng/m ³	µg/m ³	µg/m ³
	TC, OPTT, OPTR at other wavelengths, OPTT at other wavelengths	ng/m ³	NA	NA
Carbon_laser	RefF_wavelength, Refl_wavelength, RefM_wavelength, TransF_wavelength, Transl_wavelength, TransM_wavelength	reading	NA	NA
Elements	Na, Mg, Al, Si, P, S, Cl, K, Ca, Ti, V, Cr, Mn, Fe, Ni, Cu, Zn, As, Pb, Se, Br, Rb, Sr, Zr	ng/m ³	µg/m ³	µg/m ³
Light absorption	fAbs	Mm ⁻¹	Mm ⁻¹	Mm ⁻¹
Composite species	OMC, NHNO, NHSO, PM ₁₀ -PM _{2.5} , Soil	NA	µg/m ³	µg/m ³

9.2 Artifacts

An artifact is defined as any increase or decrease of material on the filter that positively or negatively biases the measurement of ambient concentration. Artifact corrections are applied to the ions, carbon, and element measurements. Artifact examples include:

Electronic documents are official. Paper copies are for reference only.

- (1) Contamination of the filter medium (positive).
- (2) Contamination acquired by contact with the cassettes or in handling (positive).
- (3) Adsorption of gases during collection that are erroneously measured as particles (positive).
- (4) Volatilization of particles during collection and in handling (negative).
- (5) Fall-off of particles during handling after collection (negative).

For the ion measurements, the artifact correction method attempts to account for the first two types of artifacts and is estimated using data from field blanks. Field blanks are handled as normal filters (loaded into cassettes and cartridges, shipped to and from the field, and left in the sampler for a week), except that no air is drawn through them. The field blanks are collected randomly at all sites on a periodic basis. When there are ≥ 50 field blanks in a month, the artifact correction is calculated for each species as the median loading measured on the field blanks. Otherwise, values from the previous month(s) are included until at least 50 field blanks are available. Artifact corrections are subtracted from each ambient concentration for the corresponding month.

For the carbon measurements, the artifact correction method attempts to account for the first three types of artifacts and is estimated using data from field blanks. The field blanks are handled as normal filters (loaded into cassettes and cartridges, shipped to and from the field, and left in the sampler for a week) except that no air is drawn through them. The field blanks are collected randomly at all sites on a periodic basis. When there are ≥ 50 field blanks in a month, the artifact correction is calculated for each species as the median loading measured on the field blanks; otherwise, values from the previous month(s) are included until at least 50 field blanks are available. Artifact corrections are subtracted from each ambient concentration for the corresponding month. For further background information and detail regarding past use of stacked filters for artifact correction and subsequent application of a correction factor, see data advisories:

http://vista.cira.colostate.edu/Improve/wp-content/uploads/2016/04/Dillner_OCArtifactAdjustmentIMPROVEOct2012.pdf and
http://vista.cira.colostate.edu/improve/Data/QA_QC/Advisory/da0032/da0032_OC_artifact.pdf

Measurements are not corrected for the two negative artifact types (volatilization and fall-off). The measured mass loadings for the higher-volatility organics may be much less than those in the atmosphere because of volatilization of particles during the remainder of the sampling or during transportation. Volatilization of nitrate and chloride from the nylon filters is assumed to be insignificant. Depending on the environmental conditions, some ammonium nitrate collected on polytetrafluoroethylene (PTFE) filters may volatilize. In those cases, fine mass on the PTFE filter may underestimate the ambient $PM_{2.5}$ mass concentrations.

For a discussion of artifact correction for element measurements, see section 9.4.4.

In the rare situations where not enough field blanks from a given lot are available according to standard practices, the analyst will take one of the following three actions:

- Perform normal data processing for the lot with fewer field blanks than are typically required; or
- Use medians of previous field blank statistics if the lot has previously been used. The field blank statistics for the current lot will include medians of the median, 95th percentile, standard deviation, and field blank count; or
- Use medians of previous field blank statistics for all lots from the same manufacturer if the lot has not previously been used. The field blank statistics for the current lot will include medians of the 95th percentile, standard deviation, and field blank count. The median will be set to zero to negate blank correction.

When the number of field blanks is less than the required minimum of 35 for PTFE filters, the field blank statistics are manually added to the database by executing the 35 required field blanks using the `get_improve_fb_set_wlot` function.

```
get_improve_fb_set_wlot( lot_number, year, month, analysis, min_samples, server = "production")
```

The `lot_number` argument is a character, while `year`, `month`, and `min_samples` are integers. The options for `analysis` argument are "xrf" and "hips". The options for the `server` are "production," "test," "development," and "sandbox."

When the number of field blanks is less than the required minimum of 50 for PTFE (mass), Nylon and/or Quartz filters, the field blank statistics are manually added to the database by executing the 50 required field blanks using the `get_improve_fb_set` function.

```
get_improve_fb_set(year, month, analysis, min_samples, server = "production")
```

The `year`, `month`, and `min_samples` are integers. The options for `analysis` argument are "ions", "carbon" and "mass". The options for the `server` are "production," "test," "development," and "sandbox."

In all of the above cases, the data will be reviewed to determine if the resulting statistics are appropriate or if another approach should be used. The analyst will add comments to the field blank sets tables in the database noting the deviation from standard procedures.

9.3 Volume

The V4 controllers automatically transmit operational data to a temporary folder in a UCD IMPROVE server daily at around 1 am local time, disregarding daylight savings. If the 1 am sync fails due to an unstable internet connection, the controller will make subsequent periodic attempts to transmit the data. The data is in CSV format. It represents

one-minute average sampling information. An automated procedure processes the sampling data into 15-minute average flow information and stores it in the module.FlowSourceDataV2 table in the UCD IMPROVE database. UCD will retain the operational data for several years before it is removed from the UCD records.

The field group at UCD can trigger data transmission if necessary. UCD utilizes backup flashcard data or log sheet records in locations with poor internet connections to obtain sampling data and average flows.

The sample volume is a product of the flow rate and the sampling duration. The sampling duration is determined using elapsed time (ET) as recorded by the sampler controller.

For the PM_{2.5} modules using the V4 controllers (1A, 2B, 3C and 5-A, B, C modules), the flow rate is determined from the measurement of static pressure across the cyclone using a pressure transducer (referred to as the CYC value). Prior to 2016, the 15-minute pressure measurements were averaged over the whole sampling period (nominally 24 hours) for calculating the average flow rate. Beginning with data for samples collected in January 2016, the average flow rate is an elapsed time-weighted average, calculated from the individual 15-minute pressure measurement. The sampler flow rate for 1A, 2B, and 3C modules is calculated using equation 351-1.

$$Q = 10^a M^b * F(elev) * \sqrt{\frac{T + 273.15}{293.15}} \quad (351-1)$$

Q = volumetric flow rate in LPM (using site-specific temperature and pressure, not STP)

a, b = calibration coefficients

M = cyclone transducer reading. If the transducer readings are taken from the controller screen, they can be used in equation 351-1 directly. If the transducer readings are taken from the flashcard file, they must be divided by 100.

$F(elev)$ = elevation factor to account for the pressure difference between sea level and site.

T = ambient temperature in degrees Celsius at the time of sampling. In this equation, T is converted to Kelvin from Celsius where 293.15 K (20 degree Celsius) is the nominal temperature.

For the PM₁₀ module (4D module), the flow rate is determined from the measurement of absolute pressure downstream of the filters near the critical orifice using a pressure transducer (referred to as the ORI value); the CYC value is not available for the 4D module. Since the pressure is measured after the filter, a decrease in measured flow rate could be indicative of a heavily loaded filter or filter clogging that is restricting the flow. The sampler flow rate is calculated using equation 351-2.

$$Q = (c + d * G) * F(elev)^2 * \sqrt{\frac{T + 273.15}{293.15}} \quad (351-2)$$

Q = volumetric flow rate

c, d = calibration coefficients

G = critical orifice transducer reading. If the transducer readings are taken from the controller screen, they can be used in equation 351-2 directly. If the transducer readings are taken from the flashcard file, they must be divided by 100.

$F(elev)$ = elevation factor to account for the pressure difference between sea level and the site.

T = ambient temperature in degrees Celsius at the time of sampling. In this equation, T is converted to Kelvin from Celsius, where 293.15 K (20 degree Celsius) is the nominal temperature.

For older V2 controllers, the ORI values can be used for the PM_{2.5} modules if the CYC reading is unreliable. All IMPROVE samplers utilize V4 controllers at present. If ORI is used, the equation 351-2 is used to calculate volume. The calibration coefficients c, d, and the critical orifice reading, G, will correspond to the specific PM_{2.5} module.

The calibration coefficients (a, b, c, and d) in equations (351-1) and (351-2) have historically been site-specific. Starting with data from samples collected January 2018, a set of universal flow constants for the V4 controller cyclone (CYC; equation 351-1) and orifice (ORI; equation 351-2). The constants are reviewed annually and updated as needed; the values are expected to vary minimally from year to year (Table 2).

Table 2. Universal flow constants for the V4 controllers.

Module	Intercept (a, c)*	Slope (b, d)*
PM _{2.5}	1.4891	0.3797
PM ₁₀	1.320	1.325

* Applied to data from 1/1/2018 onward.

9.4 Concentration, Uncertainty, and Method Detection Limit

The calculations described in this section are performed in R using the R function listed at the beginning of section 9.

The concentration is calculated using equation 351-3, where the mass of material on the filter is equal to the difference between the mass measured on the sample and the mass on

Section 508 Compliant ☐ Yes ☒ No

Data Processing
UCD TI #351B, Version 1.4
January 7, 2025
Page 15 of 33

the unused filter. For gravimetric analysis, the mass on the unused filter is determined from the pre-weight of individual PTFE filters. For the measurement of ions and carbon, the mass on the unused filter is determined from the median of field blank loadings. For calculation of element concentrations, see section 9.4.4.

$$C = \frac{A - B}{V} \quad (351-3)$$

C = ambient concentration (ng/m³)

A = mass measured on sample (ng/filter or ng/cm²)

B = artifact mass (ng/filter or ng/cm²) = pre-weight or monthly median of ion or carbon field blank mass loading

V = sample air volume (m³) = Q * Elapsed Time

Q = volumetric flow rate in LPM

The uncertainty and MDLs are reported with each concentration. Uncertainties and MDLs for ions, carbon, and elements are calculated using the following equations:

$$\sigma_{(C)} = 1000 \frac{ng}{\mu g} * \frac{\sqrt{(0.608 * \text{Max}(P95 - B, \text{mdl}_{\text{analytical}}))^2 + (f * (A - B))^2}}{V} \quad (351-4)$$

$$\text{mdl}_{(C)} = 1000 \frac{ng}{\mu g} * \frac{\text{Max}(P95 - B, \text{mdl}_{\text{analytical}})}{V} \quad (351-5)$$

Where,

V = Module sample air volume (m³)

P95 = 95th percentile of field blank measurements in µg/filter

B = artifact mass (ng/filter or ng/cm²) = pre-weight or monthly median of ion or carbon field blank mass loading

mdl_{analytical} = analytical MDL reported from the analytical laboratory. The analytical MDL is considered the ‘floor value’ and is used as the reported MDL in the event that the median value of the field blanks is lower than the respective analytical MDL.

f = fractional uncertainty. This term results from various sources of proportional uncertainties, such as analytical calibration and flow rate measurements. Beginning with data from samples collected January 2018, fractional uncertainties (f) are determined using the most recent two years of data from collocated measurements

(351-6 and 351-7). If the count of collocated pairs over the two-year period is less than 60, a value of 0.25 is adopted as f .

$0.608 = 1 / 1.645$; used to estimate the one-sigma uncertainty at zero concentration from the MDL that is set at the 95th percentile, where 1.645 is the critical value for sigma in a one-tailed test for 5% significance.

$$srd = \frac{(Collo - Routine) / \sqrt{2}}{(Collo + Routine) / 2} \quad (351-6)$$

$$f = \frac{(84th\ percentile\ of\ srd) - (16th\ percentile\ of\ srd)}{2} \quad (351-7)$$

The *improve_fracUnc* function is run using the *crocker* R package to calculate and post a new set of fractional uncertainties as well as to replace older sets, when necessary. The date range specified must be for a two-year period prior to the current year of data to be processed. The function can also be used for other purposes where the user can specify any time period of interest.

```
improve_fracUnc(startdate, enddate, effective date, server = "production",
AnalysisQcCode = 1, comment = "", replacingId = NULL, replacingQcCode = NULL)
```

For example, processing the 2019 concentration data should use the fractional uncertainties (f) calculated from 1/1/2017 through 12/31/2018 data. The function *improve_fracUnc* calculates and directly imports fractional uncertainty into database tables, *Improve_2.1.analysis.UncertaintySets* and *Improve_2.1.analysis.Uncertainties*.

```
improve_fracUnc(startdate = "2017-01-01", enddate = "2018-12-31", effective date =
"2019-01-01", server = 'production', comment = "New set to be applied beginning with
2019 data")
```

For further details, refer to the function help file in R.

9.4.1 PM_{2.5} and PM₁₀ Mass (1A and 4D Modules)

PM_{2.5} mass is measured gravimetrically on the PTFE filter from the 1A Module. PM₁₀ mass is measured gravimetrically on the PTFE filter from the 4D Module. The pre-and post-weights (as milligrams per filter) are stored in the *grav.SampleAnalysis* table in the UCD IMPROVE database.

The constant analytical uncertainty, σ_a , in equation 351-4 is equal to 5 μg for all filters. The mass concentration (C_{Mass}), uncertainty (σ_{Mass}), and MDL (mdl_{Mass}) in nanograms per cubic meter are calculated using the following equations:

Section 508 Compliant ☐ Yes ☒ NoData Processing
UCD TI #351B, Version 1.4
January 7, 2025
Page 17 of 33

$$C_{Mass} = 10^6 \frac{ng}{mg} * \left(\frac{Postweight - preweight}{V} \right) \quad (351-8)$$

$$\sigma_{Mass} = 1000 \frac{ng}{\mu g} * \frac{\sqrt{(0.608 * Max(P95, mdl_{analytical}))^2 + (f * (postweight - preweight))^2}}{V} \quad (351-9)$$

$$mdl_{Mass} = 1000 \frac{ng}{\mu g} * \frac{Max(P95, mdl_{analytical})}{V} \quad (351-10)$$

Where,

postweight = mass of filter after sampling*preweight* = mass of filter before samplingV = A-Module sample air volume (m³)P95 = 95th percentile of field blank measurements in µg/filter

mdl_{analytical} = analytical MDL reported from the analytical laboratory (Table 3). The analytical MDL is considered the ‘floor value’ and is used as the reported MDL in the event that the median value of the field blanks is lower than the respective analytical MDL.

f = fractional uncertainty (Table 4). The values for 1995-2016 have been removed from the table to save space. Those values are available in version 1.1 of this document (IMPROVE_TI_351B_v1.1_Data Processing).

0.608 = 1 / 1.645; used to estimate the one-sigma uncertainty at zero concentration from the MDL that is set at the 95th percentile, where 1.645 is the critical value for sigma in a one-tailed test for 5% significance.

Table 3. Analytical method detection limits (MDL) in µg/filter for mass.

Species	Analytical MDLs used for data			2023
	2020	2021	2022	
PM _{2.5}	10	10	10	10
PM ₁₀	10	10	10	10

Table 4. Fractional uncertainty for the mass.

Species	f reported for data						
	2017	2018	2019	2020	2021	2022	2023
PM _{2.5}	0.03	0.04	0.04	0.04	0.05	0.04	0.04
PM ₁₀	0.07	0.07	0.08	0.07	0.05	0.05	0.04

9.4.2 Ions (2B Module)

Ions are measured by ion chromatography using the nylon filter from the 2B Module. Ions data (as micrograms per filter) are stored in the *ions.MassLoadings* table in the UCD IMPROVE database.

The concentration (C_{ion}), uncertainty (σ_{ion}), and MDL (mdl_{ion}) in nanograms per cubic meter are calculated for the ion species using the following equations; however, for nitrite, when the concentration is less than or equal to zero, uncertainty is reported as zero:

$$C_{ion} = 1000 \frac{ng}{\mu g} * \frac{(A_{ion} - B_{ion})}{V_{B module}} \quad (351-11)$$

$$\sigma_{ion} = 1000 \frac{ng}{\mu g} * \frac{\sqrt{(0.608 * \text{Max}(P95 - B_{ion}, mdl_{analytical}))^2 + (f * (A_{ion} - B_{ion}))^2}}{V_{B Module}} \quad (351-12)$$

$$mdl_{ion} = 1000 \frac{ng}{\mu g} * \frac{\text{Max}(P95 - B_{ion}, mdl_{analytical})}{V_{B Module}} \quad (351-13)$$

Where,

A_{ion} = ambient mass loading in $\mu g/\text{filter}$

B_{ion} = median of the field blank mass loading in $\mu g/\text{filter}$ when there are ≥ 50 field blanks in a month; In instances where the required minimum number of field blanks is unavailable, data from three months prior or three months following the specified period will be utilized to calculate the median. Should the minimum requirement not be satisfied, the field blank statistics is calculated as per the formula detailed in section 9.2..

$V_{B module}$ = B-Module sample air volume (m^3)

P95 = 95th percentile of field blank measurements in $\mu g/\text{filter}$

$mdl_{analytical}$ = analytical MDL in $\mu g/\text{filter}$ reported from the analytical laboratory (Table 5). The analytical MDL is considered the ‘floor value’ and is used as the reported MDL in the event that the median value of the field blanks is lower than the respective analytical MDL. The values for 2006-2019 have been removed from the table to save space. Those values are available in version 1.1 of this document (IMPROVE_TI_351B_v1.1_Data Processing).

f = fractional uncertainty (Table 6). The values for 2005-2016 have been removed from the table to save space. Those values are available in version 1.1 of this document (IMPROVE_TI_351B_v1.1_Data Processing).

$0.608 = 1 / 1.645$; used to estimate the one-sigma uncertainty at zero concentration from the MDL that is set at the 95th percentile, where 1.645 is the critical value for sigma in a one-tailed test for 5% significance.

Table 5. Analytical method detection limits (MDL) in $\mu\text{g}/\text{filter}$ for the ions species.

Species	Analytical MDLs used for data			
	2020	2021	2022	2023
Chloride (Cl^-)	0.1	0.1	0.1	0.12
Nitrite (NO_2^-)	0.2	0.2	0.2	0.2
Nitrate (NO_3^-)	0.16	0.16	0.16	0.16
Sulfate ($\text{SO}_4^{=}$)	0.22	0.22	0.22	0.22

Table 6. Fractional uncertainty for ions.

Species	f reported for data						
	2017	2018	2019	2020	2021	2022	2023
Chloride (Cl^-)	0.08	0.08	0.09	0.1	0.09	0.1	0.1
Nitrite (NO_2^-)	0.25	0.25	0.25	0.25	0.25	0.25	0.25
Nitrate (NO_3^-)	0.03	0.04	0.04	0.04	0.04	0.03	0.04
Sulfate ($\text{SO}_4^{=}$)	0.02	0.02	0.03	0.02	0.01	0.02	0.02

9.4.3 Carbon (3C Module)

Carbon is measured by thermal/optical reflectance (TOR) and thermal-optical transmittance (TOT) using the quartz filter from the 3C Module. The seven carbon fractions (OC1-OC4, EC1-EC3) and organic pyrolyzed carbon (OP) are recorded in micrograms per filter and stored in the *dricarbon.MassLoadings* table in the UCD IMPROVE database. For the carbon fractions, the primary factors that determine the fractional uncertainty are the homogeneity of the sample deposit and the accuracy of the temperature set point in each stage. For OP, the primary factors that determine the fractional uncertainty are the laser signal stability and the accuracy of the split point placement.

The TOR elemental carbon (ECTR) component is assumed to be all carbon evolved at 580°C and above after the laser indicates that reflectance has returned to the initial value, minus the OP fraction. The TOR organic carbon (OCTR) component is assumed to be all carbon evolved at 580°C and below in a pure helium environment, plus the OP fraction. The total carbon (TC) is the sum of OCTR and ECTR.

Section 508 Compliant ☐ Yes ☒ NoData Processing
UCD TI #351B, Version 1.4
January 7, 2025
Page 20 of 33

The concentration, uncertainty, and MDL in nanograms per cubic meter for the carbon species (OC1, OC2, OC3, OC4, OPTR, OPTT, EC1, EC2, EC3, as well as OCTR, ECTR, TC) are calculated using the following equations:

$$C = 1000 \frac{ng}{\mu g} * \frac{(A_{carbon} - B_{carbon})}{V_{Cmodule}} \quad (351-14)$$

$$\sigma_{Carbon} = 1000 \frac{ng}{\mu g} * \frac{\sqrt{(0.608 * \text{Max}(P95 - B_{carbon}, \text{mdl}_{analytical}))^2 + (f * (A_{carbon} - B_{carbon}))^2}}{V_{CModule}} \quad (351-15)$$

$$\text{mdl}_{Carbon} = 1000 \frac{ng}{\mu g} * \frac{\text{Max}(P95 - B_{carbon}, \text{mdl}_{analytical})}{V_{CModule}} \quad (351-16)$$

Where,

A_{carbon} = ambient mass loading in $\mu\text{g}/\text{filter}$

B_{carbon} = median of the field blank mass loading in $\mu\text{g}/\text{filter}$ when there are ≥ 50 field blanks in that month, In instances where the required minimum number of field blanks is unavailable, data from three months prior or three months following the specified period will be utilized to calculate the median. Should the minimum requirement not be satisfied, the field blank statistics is calculated as per the formula detailed in section 9.2. $V_{CModule}$ = C-Module sample air volume (m^3)

$P95 = 95^{\text{th}}$ percentile of field blank measurements in $\mu\text{g}/\text{filter}$

$\text{mdl}_{analytical}$ = analytical MDL in $\mu\text{g}/\text{filter}$ reported from the analytical laboratory (Table 7). The analytical MDL is considered the ‘floor value’ and is used as the reported MDL in the event that the median value of the field blanks is lower than the respective analytical MDL. The values for 2006-2019 have been removed from the table to save space. Those values are available in version 1.1 of this document (IMPROVE_TI_351B_v1.1_Data Processing).

f = fractional uncertainty (Table 8). The values for 2005-2016 have been removed from the table to save space. Those values are available in version 1.1 of this document (IMPROVE_TI_351B_v1.1_Data Processing).

$0.608 = 1 / 1.645$; used to estimate the one-sigma uncertainty at zero concentration from the MDL that is set at the 95th percentile, where 1.645 is the critical value for sigma in a one-tailed test for 5% significance.

Section 508 Compliant ☐ Yes ☒ NoData Processing
UCD TI #351B, Version 1.4
January 7, 2025
Page 21 of 33

Table 7. Analytical method detection limits (MDL) in µg/filter for the carbon species.

Species	Analytical MDLs used for data			
	2020	2021	2022	2023
OC1	0.03	0.03	0.02	0.01
OC2	0.06	0.04	0.05	0.02
OC3	0.18	0.16	0.16	0.13
OC4	0.12	0.12	0.13	0.11
OPTR	0.12	0.14	0.15	0.09
OPTR at 405 nm	0.03	0.05	0.06	0.03
OPTR at 445 nm	0.06	0.11	0.13	0.03
OPTR at 532 nm	0.08	0.15	0.11	0.02
OPTR at 780 nm	0.08	0.12	0.11	0.1
OPTR at 808 nm	0.06	0.06	0.06	0.05
OPTR at 980 nm	0.12	0.04	0.06	0.08
OPTT	0.22	0.22	0.24	0.29
OPTT at 405 nm	0.18	0.26	0.25	0.28
OPTT at 445 nm	0.21	0.22	0.24	0.26
OPTT at 532 nm	0.19	0.24	0.24	0.27
OPTT at 780 nm	0.2	0.19	0.21	0.28
OPTT at 808 nm	0.19	0.2	0.24	0.29
OPTT at 980 nm	0.15	0.21	0.22	0.26
EC1	0.07	0.06	0.05	0.04
EC2	0.22	0.27	0.23	0.28
EC3	0.01	0.01	0.01	0.01
ECTR	0.23	0.27	0.22	0.29
OCTR	0.31	0.33	0.35	0.26
TC	0.43	0.45	0.44	0.41

Section 508 Compliant ☐ Yes ☒ NoData Processing
UCD TI #351B, Version 1.4
January 7, 2025
Page 22 of 33

Table 8. Fractional uncertainty for the carbon species.

Species	f reported for data					
	2017	2018	2019	2021	2022	2023
OC1	0.27	0.23	0.24	0.17	0.23	0.28
OC2	0.13	0.11	0.1	0.09	0.11	0.13
OC3	0.13	0.13	0.11	0.09	0.11	0.12
OC4	0.13	0.13	0.14	0.16	0.16	0.16
OPTR	0.16	0.2	0.21	0.19	0.21	0.22
OPTR at 405 nm	0.18	0.18	0.19	0.19	0.24	0.24
OPTR at 445 nm	0.17	0.17	0.18	0.18	0.22	0.22
OPTR at 532 nm	0.2	0.21	0.21	0.2	0.25	0.23
OPTR at 780 nm	0.19	0.21	0.22	0.22	0.24	0.23
OPTR at 808 nm	0.19	0.2	0.21	0.23	0.26	0.24
OPTR at 980 nm	0.21	0.23	0.25	0.24	0.27	0.26
OPTT	0.12	0.14	0.15	0.13	0.15	0.17
OPTT at 405 nm	0.13	0.13	0.14	0.12	0.15	0.16
OPTT at 445 nm	0.13	0.13	0.15	0.13	0.15	0.16
OPTT at 532 nm	0.13	0.14	0.15	0.13	0.15	0.17
OPTT at 780 nm	0.13	0.14	0.16	0.14	0.16	0.17
OPTT at 808 nm	0.13	0.15	0.16	0.14	0.17	0.16
OPTT at 980 nm	0.14	0.16	0.17	0.14	0.16	0.17
EC1	0.1	0.11	0.11	0.12	0.14	0.15
EC2	0.18	0.19	0.21	0.22	0.25	0.27
EC3	0.25	0.25	0.25	0.25	0.25	0.25
ECTR	0.14	0.14	0.13	0.14	0.14	0.16
OCTR	0.09	0.08	0.07	0.06	0.07	0.08
TC	0.08	0.07	0.07	0.06	0.07	0.08

9.4.4 Elements (1A Module)

Elements are measured using X-ray fluorescence (XRF; PANalytical Epsilon 5) using the PTFE filters from the 1A Module.

Electronic documents are official. Paper copies are for reference only.

The PANalytical XRF instruments report the elements in terms of counts per mV per second, which is converted into areal densities using element calibration factors. These calibration factors are stored in the UCD IMPROVE database in the xrf.Calibrations table. The device counts are recorded in the xrf.MassLoadings table. Blank subtraction is performed on the XRF measurements by subtracting the median field blank count from the same filter lot as that of the sample filters. The field blank correction is specific to each filter lot and since the number of field blanks from a filter lot used in each month may not be statistically sufficient, a minimum of 35 field blanks are required before the median can be calculated. Field blank selection is therefore expanded to include field blanks from previous month(s) until at least 35 field blanks are found. The selected 35 field blanks are used to calculate batch and filter lot-specific blank correction. Areal uncertainty (U_{element}) is calculated as,

$$U_{\text{element}} = 1000 \frac{\text{ng}}{\mu\text{g}} * \sqrt{(0.608 * \text{Max}((P95 - B_e), \text{mdl}_{\text{analytical}}))^2 + (f * (A_e - B_e))^2} \quad (351-17)$$

A_e = areal density calculated for the element measured by XRF.

B_e = median areal density of the field blank measured by XRF; ≥ 35 field blanks from before the determination date. In instances where the required minimum number of field blanks is unavailable, data from three months prior or three months following the specified period will be utilized to calculate the median. Should the minimum requirement not be satisfied, the field blank statistics is calculated as per the formula detailed in section 9.2.

$P95$ = 95th percentile of field blank measured by XRF.

$\text{mdl}_{\text{analytical}}$ = analytical MDL in $\mu\text{g}/\text{cm}^2$ reported from the analytical laboratory (Table 9). The analytical MDL is considered the ‘floor value’ and is used as the reported MDL in the event that the median value of the field blanks is lower than the respective analytical MDL. The values for 2006-2019 have been removed from the table to save space. Those values are available in version 1.1 of this document (IMPROVE_TI_351B_v1.1_Data Processing).

f = fractional uncertainty (Table 10). The values for 2005-2016 have been removed from the table to save space. Those values are available in version 1.1 of this document (IMPROVE_TI_351B_v1.1_Data Processing).

$0.608 = 1 / 1.645$; used to estimate the one-sigma uncertainty at zero concentration from the MDL that is set at the 95th percentile, where 1.645 is the critical value for sigma in a one-tailed test for 5% significance.

Section 508 Compliant ☐ Yes ☒ NoData Processing
UCD TI #351B, Version 1.4
January 7, 2025
Page 24 of 33Table 9. Analytical method detection limits (MDL) in $\mu\text{g}/\text{cm}^2$ for the elemental species.

Species	Analytical MDLs Used for Data			
	2020	2021	2022	2023
Al	0.011	0.011	0.013	0.013
As	0.002	0.002	0.001	0.001
Br	0.001	0.001	0.001	0.001
Ca	0.003	0.003	0.009	0.009
Cl	0.002	0.002	0.002	0.002
Cr	0.001	0.001	0.001	0.001
Cu	0.001	0.001	0.004	0.004
Fe	0.003	0.003	0.01	0.01
K	0.001	0.001	0.003	0.003
Mg	0.02	0.02	0.035	0.032
Mn	0.002	0.002	0.002	0.002
Na	0.046	0.046	0.074	0.073
Ni	0.001	0.001	0.001	0.001
P	0.002	0.002	0.002	0.002
Pb	0.003	0.003	0.005	0.005
Rb	0.002	0.002	0.002	0.002
S	0.001	0.001	0.003	0.003
Se	0.001	0.001	0.002	0.002
Si	0.005	0.005	0.014	0.014
Sr	0.001	0.001	0.002	0.002
Ti	0.001	0.001	0.002	0.002
V	0.001	0.001	0.001	0.001
Zn	0.002	0.002	0.005	0.005
Zr	0.007	0.007	0.011	0.01

Table 10. Fractional uncertainty for the elemental species.

Species	f Reported for Data						2023
	2017	2018	2019	2020	2021	2022	
Al	0.08	0.08	0.09	0.1	0.12	0.12	0.12
As	0.21	0.25	0.25	0.25	0.25	0.25	0.25
Br	0.11	0.1	0.09	0.09	0.09	0.09	0.09
Ca	0.07	0.06	0.07	0.09	0.1	0.1	0.1
Cl	0.18	0.14	0.14	0.16	0.18	0.18	0.16
Cr	0.17	0.15	0.17	0.16	0.15	0.15	0.16
Cu	0.11	0.13	0.1	0.1	0.11	0.1	0.1
Fe	0.06	0.05	0.06	0.08	0.09	0.09	0.09
K	0.05	0.03	0.04	0.05	0.06	0.05	0.06
Mg	0.16	0.15	0.15	0.17	0.2	0.19	0.19
Mn	0.13	0.14	0.13	0.13	0.13	0.13	0.14
Na	0.15	0.14	0.14	0.15	0.16	0.16	0.18
Ni	0.16	0.13	0.14	0.18	0.14	0.14	0.17
P	0.33	0.27	0.3	0.3	0.35	0.35	0.33
Pb	0.13	0.14	0.15	0.25	0.19	0.19	0.15
Rb	0.25	0.25	0.25	0.25	0.25	0.25	0.25
S	0.03	0.02	0.03	0.03	0.03	0.03	0.03
Se	0.12	0.25	0.25	0.25	0.25	0.25	0.25
Si	0.07	0.06	0.07	0.09	0.1	0.1	0.1
Sr	0.14	0.13	0.14	0.14	0.15	0.14	0.15
Ti	0.09	0.09	0.09	0.11	0.12	0.11	0.11
V	0.14	0.17	0.17	0.12	0.16	0.2	0.16
Zn	0.08	0.08	0.08	0.08	0.09	0.07	0.06
Zr	0.25	0.25	0.25	0.25	0.25	0.23	0.25

Areal densities, areal uncertainty, and areal MDL (in units of mass/area) are calculated during processing of XRF results. The concentration (C_{element}), uncertainty (σ_{element}), and MDL ($\text{mdl}_{\text{element}}$) in nanograms per cubic meter for the element species are calculated using the following equations:

$$C_{\text{element}} = 1000 \frac{\text{ng}}{\mu\text{g}} * \frac{(A_e - B_e) * (\text{Deposit area})}{V} \quad (351-18)$$

$$\sigma_{\text{element}} = \frac{(U_e) * (\text{Deposit area})}{V} \quad (351-19)$$

$$mdl_{element} = 1000 \frac{ng}{\mu g} * \frac{Max((P95 - B_e), mdl_{analytical}) * (Deposit\ area)}{V} \quad (351-20)$$

Where,

A_e = areal density calculated for the element measured by XRF.

B_e = median areal density of the field blank measured by XRF; ≥ 35 field blanks from before the determination date.

Deposit area = area of sample deposit on the filter (cm^2), determined from the filter holder or mask size (approximately 20 mm).

U_e = areal uncertainty reported for the element measured by XRF.

P95 = 95th percentile of field blank measured by XRF.

$mdl_{analytical}$ = analytical MDL reported from the analytical laboratory. The analytical MDL is considered the 'floor value' and is used as the reported MDL in the event that the median value of the field blanks is lower than the respective analytical MDL.

V = 1A Module sample air volume (m^3).

9.4.5 Laser Absorption (1A Module)

Optical absorption is measured by a hybrid integrating plate and sphere (HIPS) system using the PTFE filter from the 1A Module. The laser absorption measurements are stored as reflectance (R) and transmittance (T) values in *hips.Results* table in the UCD IMPROVE database.

Results from the HIPS measurement are reported as filter absorption coefficient ($fAbs$) in units of Mm^{-1} , calculated from R and T. The concentration ($fAbs$), uncertainty (σ_{fAbs}), and MDL (mdl_{fAbs}) are calculated using the following equations:

$$fAbs = 100 * \frac{\tau_{633} * (Deposit\ Area)}{V_{A\ Module}} \quad (351-21)$$

Where,

$V_{A\ Module}$ = 1A Module sample air volume (m^3)

Deposit area = area of sample deposit on the filter (cm^2), determined from the filter holder or mask size (approximately 20 mm).

$$\tau_{633} = \log \left(Max \left(\frac{intercept + (slope * reflectance)}{transmittance}, 0.1 \right) \right)$$

Section 508 Compliant ☐ Yes ☒ No

$$\sigma f_{Abs} = 100 * \frac{\sqrt{\left(\frac{1}{1.65} * Max(P95, mdl_{analytical})\right)^2 + (f_{unitless} * \tau_{633})^2 * (Deposit Area)}}{V_{A Module}} \tag{351-22}$$

Where,
P95 = 95th percentile of field blank measurements.

mdl_{analytical} = analytical MDL reported from the analytical laboratory (Table 11). The analytical MDL is considered the ‘floor value’ and is used as the reported MDL in the event that the median value of the field blanks is lower than the respective analytical MDL.

V_{A Module} = 1A Module sample air volume (m³)

f_{unitless} = unitless fractional uncertainty calculated from fractional uncertainty (Table 12) and nominal sample volume. The values for 1995-2016 have been removed to save space. Those values are available in version 1.1 of this document (IMPROVE_TI_351B_v1.1_Data Processing).

$$mdl_{fAbs} = 100 * \frac{Max(P95, mdl_{analytical}) * (Deposit Area)}{V_{A Module}} \tag{351-23}$$

Where,
P95 = 95th percentile of field blank measurements.

mdl_{analytical} = analytical MDL reported from the analytical laboratory (Table 11). The analytical MDL is considered the ‘floor value’ and is used as the reported MDL in the event that the median value of the field blanks is lower than the respective analytical MDL.

V_{A Module} = 1A Module sample air volume (m³)

Deposit area = area of sample deposit on the filter (cm²), determined from the filter holder or mask size (approximately 20 mm).

Table 11. Analytical method detection limit for τ₆₃₃ (unitless).

Species	Analytical MDLs used for data			
	2020	2021	2022	2023
tau_633	0.009	0.009	0.009	0.009

Table 12. Fractional uncertainty for the laser absorption data.

Species	f reported for data						
	2017	2018	2019	2020	2021	2022	2023
fAbs	0.06	0.06	0.05	0.06	0.07	0.06	0.07

In the rare situations where not enough field blanks from a given lot are available according to standard practices, see section 9.2 for guidelines on calculating field blank statistics.

9.5 Equations of Composite Variables

The following composite variables are combinations of the measured concentrations and are used in the Level 2 validation procedures described in *UCD IMPROVE #351C: Data Validation*. For the composite variables, concentration is determined along with the uncertainty and MDL. The uncertainty calculations assume that the component concentrations are independent, and the multiplicative factors have no uncertainty. The independence assumption is not strictly valid for many composites because of common factors, such as volume. However, the effect on the overall uncertainty is too small to warrant more complicated calculations.

9.5.1 Sulfate (3× sulfur from XRF) and Ammonium Sulfate (NH₄SO₄)

Sulfur is predominantly present as sulfate in the atmosphere. To compare the sulfur by XRF and the sulfate by ion chromatography, the XRF concentration is multiplied by the ratio of sulfate to sulfur atomic mass (96.06/32.06 = 3.0). This composite is labeled S3 in the data validation plots.

The sulfate is generally present as ammonium sulfate, (NH₄)₂SO₄, although it can be present as ammonium bisulfate, (NH₄)HSO₄, sulfuric acid, H₂SO₄, gypsum, CaSO₄·2H₂O, and, in marine areas, as sodium sulfate, Na₂SO₄. In many cases, the particle will include associated water, this is omitted from the calculation. In order to simplify the calculation, all sulfur is assumed to be present as ammonium sulfate. The concentrations (NH₄SO₄ and S3), uncertainties ($\sigma_{\text{NH}_4\text{SO}_4}$ and σ_{S3}), and MDLs ($\text{mdl}_{\text{NH}_4\text{SO}_4}$ and mdl_{S3}) for ammonium sulfate (NH₄SO₄) and sulfate calculated from XRF sulfur (S3) are calculated using the following equations:

$$\text{NH}_4\text{SO}_4 = 4.125 * S$$

$$S3 = 3 * S \quad (351-24)$$

$$\sigma_{\text{NH}_4\text{SO}_4} = 4.125 * \sigma(S)$$

$$\sigma_{\text{S3}} = 3 * \sigma(S) \quad (351-25)$$

$$mdl(NHSO) = 4.125 * mdl(S)$$

$$mdl(S3) = 3 * mdl(S) \quad (351-26)$$

For ammonium bisulfate, sulfuric acid, and sodium sulfate the factors are 3.59, 3.06, and 4.43, respectively. In the first two cases, the actual dry mass associated with sulfate is less than NHSO, and in the third case, more.

9.5.2 Ammonium Nitrate (NHNO)

This composite is the total dry concentration associated with nitrate, assuming 100% neutralization by ammonium. The concentrations (NHNO), uncertainties σ_{NHNO} , and MDLs (mdl_{NHNO}) are calculated using the following equations:

$$NHNO = 1.29 * NO_3^- \quad (351-27)$$

$$\sigma_{NHNO} = 1.29 * \sigma(NO_3^-) \quad (351-28)$$

$$mdl(NHNO) = 1.29 * mdl(NO_3^-) \quad (351-29)$$

9.5.3 Soil

The soil component consists of the sum of the predominantly soil elements measured by XRF, multiplied by a coefficient to account for oxygen for the normal oxide forms (Al_2O_3 , SiO_2 , CaO , K_2O , FeO , Fe_2O_3 , TiO_2), and augmented by a factor to account for other compounds not included in the calculation, such as MgO , Na_2O , water, and CO_2 . The following assumptions are made:

- Fe is split equally between FeO (oxide factor of 1.29) and Fe_2O_3 (oxide factor of 1.43), giving an overall Fe oxide factor of 1.36.
- Fine K has a non-soil component from smoke. Based on the K/Fe ratio for average sediment (*Handbook of Chemistry and Physics*), $0.6 * Fe$ is used as a surrogate for soil K. The oxide factor for K $\left(K_2O, \frac{39.1 * 2 + 16.0 \text{ g/mol}}{39.1 * 2 \text{ g/mol}} = 1.2 \right)$ is added for a total Fe factor of $0.72 * Fe$ ($0.6 * 1.2$) for the potassium oxide in soil. This increases the factor for Fe from 1.36 to 2.08.
- The oxide forms of the soil elements account for 86% of average sediment; in order to obtain the total mass associated with soil, the final factors are divided by 0.86 (*Handbook of Chemistry and Physics*). The concentrations, uncertainties, and MDLs are calculated using the following equations:

$$SOIL = 2.2 * \max(Al,0) + 2.49 * \max(Si,0) + 1.63 * \max(Ca,0) + 2.42 * \max(Fe,0) + 1.94 * \max(Ti,0) \quad (351-30)$$

$$\sigma(SOIL) = \sqrt{(2.2 * \max(\sigma(Al),0))^2 + (2.49 * \max(\sigma(Si),0))^2 + (1.63 * \max(\sigma(Ca),0))^2 + (2.42 * \max(\sigma(Fe),0))^2 + (1.94 * \max(\sigma(Ti),0))^2} \quad (351-31)$$

$$mdl(SOIL) = 0 \quad (351-32)$$

The soil variable is calculated for all valid XRF analyses.

9.5.4 Non-Soil Potassium (KNON)

Non-soil potassium is the measured fine potassium minus the soil potassium estimated from iron. Non-soil potassium is a qualitative tracer of smoke. However, the ratio of potassium/smoke mass may change as the aerosol ages. Particulate smoke potassium may be produced by the transformation of volatilized potassium and appears to be in a smaller size range than most smoke mass. Close to the smoke source, the particulate potassium may not have time to form. For long-range transport, most other smoke mass may settle out more than potassium mass. The concentrations, uncertainties, and MDLs are calculated using the following equations:

$$KNON = (K - 0.6 * Fe) \quad (351-33)$$

$$\sigma(KNON) = \sqrt{\sigma^2(K) + [0.6 * \sigma(Fe)]^2} \quad (351-34)$$

$$mdl(KNON) = 0 \quad (351-35)$$

The soil factor of 0.6 may vary slightly with the site; this will produce a small positive or negative offset for baseline values when no smoke is present. Therefore, negative values are retained. KNON is calculated for all valid XRF analyses. If a concentration is less than the MDL, the concentration and uncertainty are assumed to be equal to the MDL.

9.5.5 Organic Carbon by Mass (OMC)

To determine the total amount of organic mass associated with organic carbon, the ratio of organic mass to organic carbon is assumed to be 1.8. The concentrations, uncertainties, and MDLs are calculated using the following equations:

$$OMC = 1.8 \times OC = 1.8 \times (O1 + O2 + O3 + O4 + OP) \quad (351-36)$$

$$\sigma_{OMC} = 1.8 \times \sigma_{OC} \quad (351-37)$$

See equation 351-14 for σ_{OC} .

$$mdl_{OMC} = 1.8 \times mdl_{OC} \quad (351-38)$$

See equation 351-15 for mdl_{OC} .

9.5.6 Black Carbon

Black carbon is estimated from the initial and final laser readings from the 3C Module quartz filter analysis. For cross-module validation, black carbon is compared to the light absorption coefficient (fAbs) measured by HIPS from the 1A Module PTFE filter.

$$BC = \frac{\ln(\text{transfinal} - \text{transinitial})}{MAC} \quad (351-39)$$

TransFinal = Final laser transmittance value of the sample

TransInitial = Initial laser transmittance value of the sample

MAC = Black carbon mass absorption cross-section and it is a constant of 23 m²/g at 632.8 nm wavelength.

9.5.7 Reconstructed Mass Using Carbon Measurements (RCMC)

Reconstructed mass is the sum of sulfate, soil, salt, elemental carbon, and organic mass. The only components not included are water and nitrate. The concentrations and uncertainties are calculated using the following equations: negative values are substituted with zero. RCMC concentration is always positive. Uncertainty is calculated as the combination of individual uncertainties. The MDL for RCMC is zero. RCMC is not valid for samples in which the elapsed time for all three PM 2.5 modules is not the same.

$$RCMC = NHSO + Soil + 1.8 \times Chloride + ECTR + OMC \quad (351-40)$$

Where,

NHSO = ammonium sulfate concentration

Soil = soil concentration

Chloride = chloride concentration as measured by IC

ECTR = elemental carbon concentration by TOR

OMC = concentration of organic mass by carbon

$$\sigma_{RCMC} = \sqrt{\sigma_{NHSO}^2 + \sigma_{Soil}^2 + (1.8\sigma_{Chloride})^2 + \sigma_{ECTR}^2 + \sigma_{OMC}^2} \quad (351-41)$$

$$mdl_{RCMC} = 0 \quad (351-42)$$

RCMC is more relevant at sites where the neutralization of sulfate may be less than 100%, at sites with high nitrate, and at marine sites.

9.5.8 Reconstructed Fine Mass (RCMN)

At sites where ammonium nitrate (NHNO) is present, adding ammonium nitrate to the RCMC can make the reconstructed mass very close to the measured value. The concentrations and uncertainties are calculated using the following equations: negative values are substituted with zero. RCMN concentration is always positive. Uncertainty is calculated as the combination of individual uncertainties. The MDL for RCMN is zero.

RCMN is not valid for samples in which the elapsed time for all three PM 2.5 modules is not the same.

$$RCMN = NHSO + NHNO + Soil + 1.8 \times Chloride + ECTR + OMC \quad (351-43)$$

Where,

NHSO = ammonium sulfate concentration

NHNO = ammonium nitrate concentration

Soil = soil concentration

Chloride = chloride concentration as measured by IC

ECTR = elemental carbon concentration by TOR

OMC = concentration of organic mass by carbon

$$\sigma_{RCMN} = \sqrt{\sigma_{NHSO}^2 + \sigma_{NHNO}^2 + \sigma_{Soil}^2 + (1.8\sigma_{Chloride})^2 + \sigma_{ECTR}^2 + \sigma_{OMC}^2} \quad (351-44)$$

$$mdl_{RCMN} = 0 \quad (351-45)$$

10. DATA PROCESSING CODE

This section describes the flow of data through the data processing code used to calculate concentration, MDL, and uncertainty for all IMPROVE parameters. Figure 1 outlines the flow of data from the sampler and analysis-specific database tables to the final results.

The function `improve_process_flow` and wrapper function `improve_calculate_and_post` from the `crocker` R package are the only functions executed directly by the Data Validator (see Section 9). The `improve_process_flow` function processes and posts flow data directly into the database. The `improve_calculate_and_post` function calls several functions sequentially to calculate first measured and then derived concentrations, as well as uncertainties and MDLs. The source code for these functions is stored in the UCD source repository.

11. QUALITY ASSURANCE AND QUALITY CONTROL

11.1 Code Development

Software for data management, processing, and validation is developed in-house by professional software engineers. The source code is managed through a code repository. Development of code changes and new applications is conducted in a development environment that parallels the production environment. Prior to deployment in production, all code changes undergo testing within a separate test environment. The testing, which is conducted by developers, managers, and users, is targeted both at the

Section 508 Compliant ☐ Yes ☒ NoData Processing
UCD TI #351B, Version 1.4
January 7, 2025
Page 33 of 33

identification of software bugs and the confirmation of valid data equivalent to the production system.

11.2 Bug Reporting

Software bugs and data management issues are tracked through JIRA tracking software. All UCD users have access to an internal JIRA website and can submit, track, and comment on bug reports.

11.3 Data Validation

Data integrity is enforced within the UCD IMPROVE database via unique primary keys and non-nullable records. Data completeness and data quality are thoroughly checked through the data validation process, as described elsewhere in this SOP.

12. REFERENCES

UCD IMPROVE SOP #351: Data Processing and Validation

UCD IMPROVE TI #351A: Data Ingest

UCD IMPROVE TI #351C: Data Validation

Lide, D.R. (Ed.) (2004). *CRC Handbook of Chemistry and Physics* (85th ed.). Boca Raton, FL: CRC Press.